General comments

The paper examines what might happen if a strong atmospheric river (AR) reached the San Gabriel Mountains soon after a major wildfire. The scenario combines the January 2025 Eaton Fire with the February 2024 AR rainfall sequence. The authors use CW3E frequency and year-to-year persistence statistics (1960-2024), together with event-based AGWA/KINEROS2 simulations of prefire and postfire conditions in the Eaton-Santa Anita and Arroyo Seco watersheds. The question is relevant to NHESS, and the wildfire-AR link is explained well, particularly in Sects. 1 and 4.3. The headline estimates are a frequency of 0.51 yr^{-1} for AR ≥ 3 events; a 38 % probability of such an event in a year following one in the previous year; around 6 % more runoff and 50 % more sediment immediately below the burn; and roughly 30 % amplification of the second flood peak.

I nevertheless think that several points need substantial revision before the paper is ready for publication:

1. The calibration does not support several of the main claims.

Only total event runoff volume was used, from a single gauge in the adjacent and less-burned Arroyo Seco watershed. The same gauge lies below a sediment-filled dam that markedly alters the hydrograph (Sect. 2.4). Yet the abstract and conclusions give considerable weight to flood-peak changes (+2.3 % for the first peak and about +30 % for the second) and to hydrograph behavior. Neither quantity was constrained during calibration.

Therefore, the peak-flow results need sensitivity or uncertainty analysis, otherwise, the claims should be scaled back.

2. There is no check of the postfire simulation against observations.

The Eaton Fire occurred in January 2025, and later storms generated documented runoff and debris-flow responses within the burn area. Even a limited qualitative comparison would provide a useful reality check on the BARC-based parameter changes.

3. The recurrence statistics are described too broadly.

The reported 38 % is the climatological probability of AR ≥ 3 in year t given AR ≥ 3 in year t−1. It is not the probability of a strong AR after wildfire.

The paper also does not show any sampling uncertainty, and the February 2024 event seems to have been weaker than AR 3 at some of the five locations.

4. The postfire parameterization is not shown.

Because the postfire signal comes entirely from these parameter changes, the adopted values and their sensitivity need to be documented. The readers need to see the adopted values and how sensitive the results are to them.

5. <u>The simulated response is modest compared with the studies cited in the Introduction.</u>

Those studies report runoff increases above 70 % and erosion increases of one or more orders of magnitude. The difference may be reasonable, but it needs to be explained.

This difference deserves discussion. Since KINEROS2 also excludes debris flows, the sediment estimates should be described as conservative lower bounds.

6. <u>The manuscript deals with hazard, not risk in the usual sense.</u>

It does not assess exposure or vulnerability, yet the title and much of the text use the term "flood risk".

7. <u>Several other problems also need attention.</u>

These include an incomplete sentence in Sect. 4.2, a wrong cross-reference, the absence of the required "Code and data availability" section, and several typographic and formatting errors listed below.

Specific comments

1. <u>Calibration vs. conclusions (Sect. 2.4, lines 179–193; Sect. 3.2, lines 219–230; Fig. 5)</u>

This is my main methodological concern. Calibration uses total event runoff volume from one event and one gauge. The gauge is below the sediment-filled Brown Mountain Dam, which the authors acknowledge has altered the hydrograph, and it lies in a watershed adjacent to the one used for the main conclusions. Neither peak flow nor timing was calibrated, although both feature prominently in the abstract and conclusions. I would therefore like to see:

a) a sensitivity or uncertainty test showing whether the relative postfire peak changes (+2.3 % and +30 %) remain stable when the calibrated parameters are varied (Ksat multiplier, CV of Ksat, capillary drive, and Manning's n);
b) the observed and simulated event volumes, the resulting bias, and an explanation of why multipliers calibrated in Arroyo Seco can be transferred to ESA; and
c) more cautious wording in the abstract and Sect. 5, since peak flow was not constrained by observations.
d) Finally, the sentence at lines 221-223 about "near-zero biases" and a "10%" runoff coefficient is unclear, please replace it with the actual values.

2. <u>Postfire validation (Sects. 2.3, 3.3, and 4.2)</u>

More than a year separates the fire from the manuscript, so the model should be compared with an observed post-Eaton storm response if possible. Evidence might come from USGS debris-flow assessments, LA County Public Works records, gauges, or agency reporting on the February-March 2025 storms. A first-order comparison would be enough to add value. If suitable observations are unavailable or not comparable, say so and treat this explicitly as a limitation in Sect. 4.2. At present, the postfire results rest entirely on uncalibrated look-up-table adjustments.

3. Postfire parameterization (Sect. 2.3.1, lines 161–164)

All postfire differences come from BARC-based changes to Manning's roughness, saturated hydraulic conductivity, and interception, but the values are not reported.

a) Add a table with the prefire and postfire values, or multipliers, for each burn-severity class and cite the source of each choice.
b) It should also be clear whether any channel-element parameters have been changed.
c) A sensitivity test over ranges supported by the literature is important because the reported +6 %, +30 %, and +50 % changes are now single deterministic outcomes.

4. Magnitude of the simulated postfire response (Sects. 3.3 and 4.2; lines 57–59 and 259–281)

The Introduction cites postfire runoff increases above 70 % and erosion increases of one or more orders of magnitude. Here, the increase is about 6 % for runoff in channels directly below the burn (about 2 % for the watershed as a whole) and roughly 50 % for sediment.

The contrast should be discussed directly. Possible reasons include saturation during an extreme AR, the dominance of moderate burn severity, aggregation of burned and unburned areas within model elements, or the difference between event-scale and annual estimates. KINEROS2 does not simulate debris flows or landslides (lines 333–336), although these processes contributed much of the sediment during the actual 2024 and 2025 events.

The abstract and conclusions should therefore identify the sediment changes as likely lower bounds.

5. Initial conditions (lines 172–173)

The initial soil saturation was set to 0.001 to represent the eight dry months before the fire. Was the same value applied to the prefire and postfire runs?

It would also be useful to show how this assumption affects the results, especially the second peak, which is attributed to storage filling at lines 241–243 and 293–295.

An AR in the year following the fire would not necessarily follow an equally long dry spell. A short sensitivity test, or at least a fuller discussion, is warranted.

6. Use of the term "risk"

The paper estimates event likelihood and hydrologic consequences such as runoff, peak flow, and sediment yield. It does not analyze exposure or vulnerability, and the likelihood and consequence components introduced at lines 89–91 are never combined.

In that sense, the analysis is of postfire flood hazard, not risk.

Either change the terminology in the title and text or define the limited qualitative meaning of “risk” and state clearly that no integrated risk metric is calculated.

7. Duration analysis (Sect. 4.1, lines 293–308; Fig. 8)

The February 2024 event is assigned a maximum duration of 36 h across the five sites, based on IVT-threshold exceedance. The rainfall sequence used in the simulation lasts nine days and has three pulses. Please clarify three points:

a) whether “[23.0, 36.0] hr” in the Fig. 8 legend is the range across the five sites;
b) whether a 36 h IVT window is the right hydrologic comparison when the flood response reflects the full multi-pulse sequence; and
c) whether the historical range is 20–120 h, as stated at line 297, or 25–120 h, as stated in the caption.

8. Figures

- Fig. 1: Replace the unclear legend heading “Burned Area Value” with “Burn severity (BARC)” and name the classes. Change “USGS_gauge” to “USGS gauge”. Coordinates or a graticule would also help.
- Fig. 3: Define “recurrence” in the caption. Consider marking the February 2024 event in the time series and enlarge the axis labels.
- Fig. 4: Correct “Arroyo-Secco” on the boxplot axis. Also explain how the comparable cumulative contributing area in ESA was chosen, including which hillslopes were used and the selection criterion.
- Fig. 5: State in the caption that panels (a) and (b) show the first and second flood peaks, respectively. Explain that outflow in mm hr^{-1} is discharge divided by contributing area. Because the prefire and postfire curve overlaps in places, a difference hydrograph or a small inset around each peak would make the comparison easier to read.
- Figs. 6–7: The class limits are irregular (for example, 0–0.22, 0.22–1.57, ...). Say how they were selected, such as by quantiles, and add “%” to the legend titles.
- All map figures, add coordinate labels and basemap attribution.

Point-by-point corrections

- Lines 203–204 state that the February 2024 event ranged from AR1 to AR3 among the five sites, depending on local IVT and duration. What was its AR-scale category at the grid cell or cells closest to the study watersheds (C/D)? If it was below AR 3 there, the paper should explain why AR ≥ 3 statistics are used to support the plausibility of an event analogous to February 2024.
- The stated “conditional recurrence” is $P(AR \geq 3 \text{ in year } t \mid AR \geq 3 \text{ in year } t-1)$, which is a persistence measure for AR climatology. It should not be presented as the probability of a strong AR following a wildfire, as it currently is in the abstract (line 11) and at lines 209–211. Please correct this distinction in the abstract, results, discussion, and conclusions.

- Please add uncertainty ranges, for example bootstrap confidence intervals, for both 0.51 yr^{-1} and 38 %. The record covers only 64 years, and the five nearby grid cells are spatially dependent rather than independent samples.
- Define what is meant by a "year" in this calculation (calendar year or water year). AR seasons cross the calendar-year boundary, so this choice can change the consecutive-year result.
- The choice of the five coastal locations also needs justification. Site A (35° N, 120.5° W) is about 200 km from the study area and has the largest frequency and recurrence values. Including it raises the multi-site averages reported in the abstract.
- Line 144 gives a 65-year CW3E record (1960–present), whereas line 205 and the Fig. 3 caption refer to 64 years (1960–2024). Use one consistent period throughout.
- Line 194: change the section heading "3 Result" to "3 Results".
- Lines 335–336: "Because such processes contributed significantly to sediment export during the 2024 AR event (Selby, 2025), sediment yield may be." is incomplete. The intended ending may be "...may be underestimated."
- Line 326: Sect. 2.5 does not exist; this appears to mean Sect. 2.4.
- Line 180: the gauge is shown in Fig. 1, not Fig. 2.
- Insert a space before parenthetical citations throughout, for example in "spread(Keeley and Syphard, 2021)", "seasons(Kumar et al., 2025)", and "networks(Goodrich et al., 2012)".
- Line 53: replace "combinedly" with "in combination".
- Lines 35–36: revise "the number of fires exceeding 10,000 ha, with several surpasses 100,000 ha have been increased" for grammar.
- Lines 37–38: change "followed by even larger event in 2021, the Dixie Fire" to "followed by an even larger event in 2021, the Dixie Fire".
- Lines 140–141: add a full stop before "The February 2024 event was classified...".
- Line 229: correct "Arryo Seco" to "Arroyo Seco". Also use one element term consistently: "Plain 63" and "Planes 11, 31..." should be "Plane" or "Hillslope", as appropriate.
- Give SI units first at lines 73 and 122–123; the current text begins with "8 inches" and "7–8 inches (≈180–200 mm)".
- Define AR once at first use. The acronym is currently redefined at several points, including lines 8, 63, 76, 95, 121, 134, and 195.
- Use "prefire/postfire" and "post-fire" consistently throughout.
- Table 1 and the references: labels such as "(Aorc, n.d.)", "(Usgs, n.d.–a)", "(Mrlc, n.d.)", "(Baer, n.d.)", and "(Gssurgo, n.d.)" should use the responsible organizations (NOAA, USGS, MRLC, USDA-NRCS, and USFS/BAER) and follow the journal style.
- References: several "last access: February 27" entries have no year; Phillips (2025) has no source or publisher; and the Kelley et al. (2025) entry may contain

a duplicated or merged author list. Please check it against the ESSD paper “State of Wildfires 2024–2025” and verify the full reference list against NHESS style.

- Fig. 8 caption and line 297: reconcile “25 to 120 hours” with “20 to 120 hours” (see comment 9c).

I do not regard these points as grounds for rejection, and I think the paper could make a useful contribution once they are addressed.

I therefore recommend **reconsideration after major revisions**.